Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 12 (4). 2022. 651-673
https://doi.org/10.14746/sslt.2022.12.4.6
http://pressto.amu.edu.pl/index.php/sslIt

WCF processing in the L2 curriculum: A look at type of WCF, type of linguistic item, and L2 performance

Ronald P. Leow
Georgetown University, Washington, DC, USA
https://orcid.org/0000-0002-2990-3540
leowr@georgetown.edu

Anne Thinglum
Georgetown University, Washington, DC, USA
https://orcid.org/0000-0002-0007-1365
amc254@georgetown.edu

Stephanie A. Leow
Arizona State University, Tempe, USA
https://orcid.org/0000-0002-9459-9346
sleow@asu.edu

Abstract

Whether type of written corrective feedback (WCF) impacts L2 learning has been investigated for decades. While many product-oriented studies report conflicting findings, the paucity of studies adopting both a process-oriented and curricular approach (e.g., Caras, 2019) underscores the call for further research on: a) the processing dimension of L2 writers' engagement with WCF in this instructed setting (Manchón & Leow, 2020), b) from an ISLA *applied* perspective (Leow, 2019a; Leow & Manchón, 2022), and c) any potential relationship with subsequent performances. Also, whether type of linguistic item (e.g., morphological vs. syntactic) plays a role in the processing dimension also warrants further probing. This preliminary quasi-experimental study explored the cognitive processes of 10 adult L2 writers with minimal previous exposure to Spanish interacting with WCF (both direct and metalinguistic) on morphological and syntactic errors. Think aloud data gathered from three

compositions written within the natural writing conditions of a foreign language curriculum were transcribed, coded for depth of processing (DoP) (Leow, 2015), and correlated with subsequent performances on the target items. The results revealed: 1) a higher DoP for metalinguistic WCF, 2) differences in processing of linguistic items, 3) similar DoP over time, and 4) a beneficial relationship between DoP and subsequent performances. Recommendations for future research underscore the importance of acknowledging variables within the instructed setting that may impact a pure effect of WCF on L2 development.

Keywords: WCF processing; type of linguistic item; ISLA *applied*; think aloud protocols; depth of processing

1. Introduction

Writing comprises an important component in many language curricula in which composing is not only promoted as an outside classroom assignment but also formally evaluated within a testing condition. Writing in the second/foreign language (L2) classroom is also closely linked to the provision of written corrective feedback (WCF), which typically forms an integral part of such written production. WCF can be viewed as "any external manipulation of L2 writers' product by the teacher or the researcher designed to minimally draw their attention to some grammatical, lexical, structural, and/or content error committed by the L2 writers" (Leow, 2020, p. 99).

The effects of type and amount of WCF on L2 development have permeated the writing strand of research for many decades (see Leow, 2020, for a recent review), while the relative paucity of studies addressing type of linguistic item (e.g., lexical, morphological, syntactic) has also been targeted as a potential variable warranting further investigation (Manchón & Leow, 2020). It can also be observed that the majority of WCF studies have approached the role of WCF from a product-oriented (versus process-oriented) and non-curricular (lacking any association with the language curriculum) perspective (Leow, 2020). A product-oriented approach relies on the product of revised compositions (after the process of revising) while a process-oriented approach gathers concurrent data to first establish the cognitive processes and strategies employed during the revision phase before addressing any potential effect of WCF on subsequent L2 development. The few studies adopting both a process-oriented and curricular approach (e.g., Caras, 2019; Coyle et al., 2018) underscore the need for further research on the processing dimension of the L2 writer's engagement with WCF in this instructed setting (Leow, 2020) and, more specifically, from an instructed second language acquisition (ISLA) *applied* perspective (Leow & Manchón,

2022; Manchón & Leow, 2020). According to Leow (2019a), *applied* ISLA comprises studies that investigate the potential effects of many variables in the instructed setting (e.g., the role of WCF in L2 development in a laboratory setting) but are usually not situated within the language curriculum. ISLA *applied* seeks to inform practical pedagogical practices by situating this same provision of WCF within the syllabus or language curriculum.

In an effort to better understand how L2 writers process type of WCF, this preliminary quasi-experimental study explored the cognitive processes and depth of processing (DoP) of adult L2 writers interacting with both direct and metalinguistic WCF on morphological and syntactic errors gathered within the natural writing conditions of a foreign language curriculum and semester-long syllabus. DoP was defined as "the relative amount of cognitive effort, level of analysis, and elaboration of intake, together with the usage of prior knowledge, hypothesis testing, and rule formation employed in decoding and encoding some grammatical or lexical item in the input" (Leow 2015, p. 204). Whether this processing varied according to type of linguistic item and over time and any relationship between DoP during rewrites and potential L2 performance were also investigated.

2. Theoretical underpinnings

There are several theoretical underpinnings in ISLA WCF literature that may account for the role of WCF during the revision phase of the L2 writing process (see Leow & Suh, 2022 for a recent review). Leow's (2020) *feedback processing framework* was selected for this study based on its cognitive explanation for the role of corrective feedback in L2 development premised on how L2 writers process such feedback. According to the framework, *feedback* on L2 writers' output comprises the L2 information to which L2 writers need to minimally pay attention in order for feedback intake to enter into their working memory. *Feedback processing* is the stage that addresses *how* L2 writers cognitively process the feedback (if at all) and is related to their current knowledge (whether accurate or inaccurate). If further processed at this stage, and regardless of levels of DoP or awareness, the framework predicts that information in the feedback allows for: 1) reinforcement of accurate prior knowledge or 2) potential restructuring of previously learned inaccurate knowledge stored in the L2 writers' *internal system*. New restructured information, irrespective of accuracy, replaces or joins the original prior knowledge in the internal system, which is then available for the *knowledge processing* stage. Of importance is the postulation that L2 writers may still retain the previous inaccurate L2 data and now possess both (accurate and inaccurate) options in the system. At the output stage, *old* (or inaccurate) *output* may indicate a potential absence or low depth of prior processing of the

WCF provided or not much confidence in the newly restructured knowledge if the feedback information was internalized. *New* or *modified output* represents L2 writers' production of the restructured L2, thus representing the L2 knowledge currently in their *internal system*. This current knowledge may be a result of item learning (a chunk of language or a simple repetition of the feedback) that may be temporary or immediate. It may also be systemized learning (available to be generated to different but similar outputs) indicating that accurate restructuring did take place. Evidence of L2 writers retaining their accurate restructuring of the feedback in future production reflects systemized learning while L2 writers reverting back to their previous inaccurate interlanguage reflects item learning. Other variables that may play a role in feedback processing include "depth of processing, levels of awareness, activation of appropriate prior knowledge, hypothesis testing, rule formulation, and/or metacognition" (Leow, 2020, p. 105).

3. Review of the literature

3.1. The revelations of WCF research

A recent review of WCF studies (Leow, 2020) reveals: 1) inconclusive findings regarding the benefits of type of WCF, 2) a higher focus on a product-oriented approach to the effect of WCF than a process-oriented one (how such WCF was processed), 3) assumptions made on WCF processing (e.g., indirect WCF promotes deeper processing while direct WCF reduces cognitive load), 4) an overall belief that type of linguistic item may play a role during the provision of WCF although there is a dearth of unfocused studies beyond English grammatical items aligned with an authentic classroom setting, 5) overall global measures to address specifically what L2 writers "learned" from WCF, and 6) experimental designs not aligned with the language curriculum and typically of a one-shot laboratory-based design offering minimal pedagogical extrapolations. However, Leow (2020) also reported an increase in WCF studies adopting 1) a *writing-to-learn* perspective of WCF provision (e.g., Manchón, 2011), 2) a process-oriented research perspective (e.g., Leow & Manchón, 2022; Manchón & Leow, 2020) evidenced by using concurrent data elicitation procedures (e.g., Caras, 2019 for online verbal reports and Manchón et al., 2020 for offline written languaging), 3) a curricular approach (e.g., Caras, 2019), and 4) an *ISLA applied* approach (Leow & Manchón, 2022; Manchón & Leow, 2020).

Given that WCF is external L2 data, whether L2 writers adequately process provided feedback or even understand it remains to be robustly established. To better understand the role WCF plays in subsequent restructuring and potential learning warrants a full investigation into how L2 writers process WCF. To this end, concurrent data gathered during the revision phase are essential to provide

relevant insights into the potential connection between the writing processes during revision and L2 learning.

3.2. Motivation for the present study

Situating the *writing-to-learn* strand of research within a process-oriented and curricular approach is important if researchers wish to extrapolate empirical findings to the instructed setting (ISLA *applied*). This approach allows researchers to address the dearth of concurrent data on the cognitive processes employed by L2 writers during the revision phases of the L2 writing process. A better understanding of how L2 writers interact with or process WCF on different types of linguistic items together with any potential relationship with subsequent L2 development will: 1) elucidate the *writing-to-learn* process and 2) allow researchers to avoid making assumptions on how L2 writers process WCF. Concurrent data will also provide insights into, for example, how L2 writers process linguistic items, perhaps based on the type (e.g., morphological vs. syntactic) and characteristics (e.g., saliency, complexity, etc.) of the error produced, and the role of DoP during the revision process. Such rich data can only lead to pedagogical implications aimed at promoting robust learning from the ISLA *applied* strand of research (Leow & Manchón, 2022; Manchón & Leow, 2020).

To this end, the research design of this study followed the one employed in Caras (2019). Data collection occurred at one level (first-semester) of the language curriculum, curriculum-based compositions were carefully designed to elicit, via prompts, target linguistic items, feedback was unfocused due to ecological validity while the two target items were selected for investigation. Furthermore, to adhere more closely to the exemplar of an ISLA *applied* research design (Leow & Manchón, 2022), several modifications were implemented: 1) compositions were written at home and not in the language laboratory, 2) no time limit was set, 3) data were gathered over three compositions with different topics during the semester, and 4) a within-subject design was employed in which participants received both types of WCF. The following research questions guided the present study:

RQ1: How do adult L2 writers process type of feedback (i.e., direct vs. metalinguistic) during the revision phases?
RQ2: Does this processing vary: (a) according to type of linguistic item (i.e., morphological vs. syntactic) and (b) over time?
RQ3: Is there any relationship between depth of processing during rewrites and L2 performances as measured on subsequent compositions, curricular tests, and final exam?

4. Research design

4.1. Participants

Ten beginning English-speaking learners of Spanish (1st of 4 semesters to fulfill a college-level language requirement) with minimal previous exposure to Spanish were drawn from the larger population of the writing research project study in progress targeting several linguistic items across different proficiency levels in the language curriculum. Participants comprised three males and seven females with an average age of 19 years. Inclusion was based on fulfilling the criteria of having minimal previous knowledge of Spanish and submitting the first three compositions assigned during the semester together with associated think aloud (TA) protocols. Together with eight other students with some previous high school knowledge of Spanish, they attended three 50-minute weekly sessions for 14 weeks (approximately 35 hours of formal exposure). Spanish was spoken almost exclusively in class after the first day and the methodology was communicative with a focus on all four skills. Participants/students were informed early in the semester about the concept of DoP and its potential contribution to any type of learning and subsequent retention.[1] They were also encouraged to follow the curricular policy of preparing the syllabus activities *before* attending class. Instruction was virtual (Zoom) due to the COVID pandemic.

4.2. Target linguistic items

Two target linguistic items, known for being problematic for early-stage English-speaking learners of Spanish as a foreign language, were selected. The morphological target linguistic item comprised the Spanish agreement between noun and adjective (e.g., *la Casa Blanca* "the House White" where the masculine adjective *blanco* changes to *blanca* to agree with the feminine noun *casa*). The syntactic target linguistic item was two levels of the *gustar* structure. As stated by Cerezo et al. (2016), referencing Housen and Simoens (2016), "Spanish gustar structures pose multiple problems for L1 English speakers from a developmental, psycholinguistic, linguistic, and pedagogical perspective, all of which add to their overall learning difficulty" (p. 273). They reported four levels of *gustar* [Level 1: *me/te gusta(n)* "I/you like (something)" > Level 2: *le/les gusta(n)* "he/she/you/they like (something)" > Level 3: *A Juan le gusta(n)* "Juan likes (something)" > Level 4: *A Juan y a Marisol les gusta(n)* "Juan and Marisol like (something)"] based on the number of steps L2 learners need to take to either process or produce each level (see Cerezo et al., 2016 for further elaboration). The current study investigated Levels 1 and 3.

[1] Being a research institution, students and the teaching staff in this department are usually kept informed of theoretically-driven and empirically-supported research findings that promote more robust learning.

4.3. Type of WCF

The two types of WCF were direct and metalinguistic, which was accompanied by a metalinguistic sheet titled *Abbreviations and correction symbols* with grammatical information (see Appendix A for a sample of the abbreviations and symbols). For direct WCF, the incorrect agreement was crossed out and replaced by either the correct adjectival or verbal bolded form as seen below:

> Agreement: Mi compañera de clase es muy ~~serio~~ *seria*.
> Gustar: Me ~~gusta~~ *gustan* mis clases este semestre.

For metalinguistic WCF, the incorrect agreement was bolded and accompanied by a symbol (e.g., AGR = Agreement) related to the metalinguistic sheet.

> Agreement: Mi compañera de clase es muy *serio* AGR.
> Gustar: Me *gusta* AGR mis clases de español.

4.4. Data

The data to address research questions 1 and 2a comprised two of the three curriculum-based compositions in which the two target linguistic items (in addition to others) were embedded in the prompts. These compositions were assigned during the regular semester based on the chapter topic, vocabulary, and grammar recently covered. The required length was approximately 300 words. Composition 1 (*Yo mismo/a y mis clases* "Myself and my classes" in Week 3 after 5 hours of formal exposure) provides a sample of the prompts designed to elicit the target linguistic items while adhering to the curricular learning objectives of the course up to that point in the semester (see Appendix B). Composition 2 (*La información personal y el horario de mis amigos* "Personal information and my friends' schedules" in Week 7 after 13 hours of formal exposure) provides an example of embedded previous target items (agreement and *gustar*) in subsequent compositions to address potential retention of WCF information (see also Appendix B). Composition 3 (*Mis actividades (pasadas y actuales) durante el año pasado y este* semestre "My activities (past and present) during the last year and this semester) was assigned in Week 12.

To address whether there was a relationship between DoP and potential L2 learning as measured within the language curriculum learning outcomes, performances were scored on the participants' responses to the WCF provided on the rewrites and also on items embedded in test 1 (administered after composition 1) and test 2 (administered after composition 2) and also in the final exam. In test 1, test 2, and the final exam, similar prompts from the previous compositions were included in the written production section.

4.5. Procedure

Participants followed the current procedure of composition writing in the language curriculum. They received a topic with several prompts to address, a set of guidelines and grammatical points to be covered, and a specific length. Subsequently, using the computer at home, they wrote and rewrote (after type of feedback was provided alternatingly, for example, metalinguistic on composition 1, direct on composition 2, etc.) their compositions. They were also given a document with instructions on the preparation (e.g., use of Voice Memos, Garage Band (iOS), Voice Recorder, Audacity, Samsung Voice Recordings, or Easy Voice Recorder) and procedure that included TA practice (see Appendix C), and instructions to upload their files to a server using their participant numbers. At the end of the semester, participants completed an online questionnaire (see Appendix D) addressing their preference of WCF, the reason for their selection, and their opinion of thinking aloud.

4.6. Coding

Participants' TA protocols were coded for DoP of the target linguistic items during the revision phase. Leow's (2015) coding scheme, premised on Leow's (2012) correlations between levels of processing and levels of awareness, was adopted for this study. More specifically, low DoP was correlated with awareness at the level of noticing (i.e., minimally mentioning the correction), medium DoP correlated with awareness at the level of reporting (i.e., making some comments), and high DoP correlated with awareness at the level of understanding of the underlying rule (Schmidt, 1990). However, Leow (2015) indicated that while high DoP may be evidenced in the TA protocol, this does not suggest that the L2 learner understood the underlying rule present in the WCF. Therefore, high DoP may be associated with either plus or minus awareness at the level of understanding. Interrater reliability among three raters was 93%, which was raised to 100% after discussion between the raters. Samples of DoP for the *gustar* structure are presented in Table 1.

Table 1 Samples of DoP for the *gustar* structure

Depth of processing	Example
+H (High DoP/+(awareness at the level of) understanding	. . . A mi madre . . . a personal . . . a mi madre . . . so she likes, so it's pleasing to her . . . I'm going to use le gustó, or she liked . . . so le gustó. [Metalinguistic]
+H (High DoP/-(awareness at the level of) understanding	. . . I forgot the "a" in the next sentence . . . so the next sentence should be A Jacqueline le gusta su clase de literatura . . . so, then I forgot the le gusta . . . [Direct]
+M (Medium DoP/+(awareness at the level of) reporting	. . . y sus amigos . . . uh, les gusta . . . I feel like I should be doing les gustan . . . but I'm not sure about that . . . I'll check that later [Metalinguistic]
+L (Low DoP)/+(awareness at the level of) noticing	. . . *se gusta su profesor* (did not address the Direct WCF provided) [Direct]

4.7. Scoring

Participants' responses to the WCF provided on the rewrites and on items embedded in Test 1, Test 2, and the Final Exam were scored 1 correct or 0 incorrect for agreement between nouns and adjectives while for *gustar* 1 (totally correct), .5 (partial, viewed globally, for example, *me gusta las clases*, *Manuel le gustan las clases* = .5), or zero (e.g., *yo gusto*) were assigned.

5. Results

To address RQ1 that sought to investigate how adult L2 writers process type of feedback (direct vs. metalinguistic) during the revision phases, the percentages of instances of DoP coded on the TA protocols for participants in each WCF condition on Table 2 revealed that, while the overall level of DoP was high for both WCF conditions, it was higher at this level for metalinguistic WCF (100% for +H) when compared to direct WCF (60% for +H). In addition, both types of WCF elicited relatively similar +H/+understanding (50% and 40% for the metalinguistic and direct, respectively).

Table 2 Depth of processing in direct and metalinguistic WCF conditions

DOP	Direct %	Metalinguistic %
+H/+understanding	40	50
+H/-understanding	20	50
+M/+reporting	20	0
+L/+noticing	20	0

The questionnaire (see Appendix D) revealed a split 50% for either type of WCF. This result was unexpected given: 1) the assumed low DoP or cognitive effort required for direct WCF and 2) the low level of proficiency that arguably lacks a sophisticated repertoire of grammatical knowledge. For Direct WCF, responses mainly focused on the ease of processing and identifying errors while critiquing the metalinguistic sheet (see Appendix A) for its complexity. Participants opting for metalinguistic WCF expressed that this type of WCF prompted them to think more and deeper in order to learn from mistakes. Interestingly, on thinking aloud, 90% reported that it allowed them to "talk" their way through the composing and rewrites and helped them to identify problem areas.

To address RQ2a, that is, whether the level of processing varied according to type of linguistic item (morphological vs. syntactic), the data revealed that participants processed both the morphological (agreement) and syntactic (Level 1 of *gustar*) linguistic items relatively equally but processed agreement substantially more than the higher level of *gustar* (Level 3). In addition, several participants mentioned

the agreement rule during both the composing and rewrite stages of the second and third compositions, indicating high awareness of this phenomenon in Spanish.

Regarding whether the level of processing WCF was maintained over the semester (RQ2b), the amount of time participants spent during the revision phases and the DoP data over the three compositions were examined. Participants spent a relatively similar average of 30, 21, and 26 minutes during the three revision phases while the DoP data indicated that they continued to report a relatively high level of processing for both linguistic items (agreement and Level 1 of *gustar*). A further analysis of the number of words and time spent on the original compositions together with their respective DoP data appeared to support participants' relatively high DoP. The range of words written from composition 1 to composition 3 was as follows: 139-436, 341-758, and 378-505 words, respectively. Almost all participants met the word requirements in all compositions, and yet the average amount of minutes spent on the original compositions was reduced substantially from 95 minutes on composition 1 to 68 and 60 minutes on the second and third compositions, respectively. The DoP data revealed several instances of reactivation of the target linguistic rules as participants composed compositions 2 and 3. For example, on the second composition: "La clase de historia es muy intensivo. Ok so here "intensivo" is adjective which is modifying "clase," so I am going to change that to "intensiva" and on the third composition: "So, the classes I like . . . so "me gustan" because I am doing multiple classes . . . ." The reduction of time in spite of longer compositions, together with the DoP data, appeared to indicate a substantial reduction of cognitive effort during composing and processing the WCF as the semester progressed. Crucially, agreement was practiced or embedded in much of the course materials and compositions, Level 1 of *gustar* was less practiced or embedded when compared to agreement and Level 3 of *gustar* even less so during the semester.

Table 3 Levels of gustar investigated in the current study

Level of gustar	Example
Level 1 *me/te gusta(n)* "I/you like (something)"	I really do not like Zoom so I can use "gustar" here . . . So I want to say I don't like Zoom so I start with "no" "no me gusta Zoom." So "no me gusta" . . . hmmm . . . "me gusta" and Zoom is doing the pleasing . . . So I'm going to conjugate it "gustar" with an "a" and not "gusto" So, "no me gusta Zoom" hmm . . . So, the classes I like . . . so "me gustan" because I am doing multiple classes, so "me gustan las clases" . . . [Participant C]
Level 3 *A Juan le gusta(n)* "Juan likes (something)"	A mi madre . . . a personal . . . a mi madre . . . so she likes, so it's pleasing to her . . . I'm going to use le gustó, or she liked . . . so le gustó. [Participant G]

To address RQ3 that sought to address whether there was a relationship between DoP and subsequent performances on the tests and final exam, data revealed that participants (70%) who demonstrated +high DoP/+understanding

showed over 95% agreement accuracy on test 1, test 2, and the final written *and* oral exams, in which agreement items were embedded. For the *gustar* structure, accurate performance required a high DoP plus awareness at the level of understanding, as reported in Table 3.

Participants (60%) demonstrating a high level of DoP, indicating awareness of the underlying rule, were 90% accurate on the tests and final exam for Level 1 while only two participants who were +H/+understanding for Level 3 were accurate on similar structures on the final exam. At the same time, lower levels of DoP revealed uncertainty with the *gustar* rule, which led to inaccurate production on the tests and final exam.

> . . . *Yo gusto* . . . so, that would be . . . change that to *me gusto* . . . I believe that's correct [+M DoP]
> Comp. #2: *se gusta su profesor* (did not address the direct WCF provided) [+L DoP]
> Comp. #3: *mi padre se gusta hacer una camineta*

Even when the WCF led to accuracy during the rewrite phase, robust internalization of the structure rule was apparently not achieved as observed in this protocol in which Participant B received metalinguistic feedback on his original composition:

> . . . Me gusta AGR mis clases de . . . no me gusta mis clases de . . . During the WCF stage:
> . . . Me uh agreement uh oh me gustan mis clases de español. No me gustan um my classes . . .

Yet he was inaccurate on the final exam, indicating a lack of processing at a level indicative of restructuring and understanding of the structural rule.

6. Discussion

The current study aimed to adopt a process-oriented ISLA research design (Leow & Manchón, 2022; Manchón & Leow, 2020) that followed an ISLA *applied* perspective (Leow, 2019a) by embedding the design within a language curriculum to gather concurrent data on L2 writers' cognitive engagement with WCF in this authentic instructed setting. The data revealed that while the overall DoP for both WCF conditions was relatively high, the provision of metalinguistic WCF led to higher DoP as compared to that observed when processing direct WCF (100% vs. 60%, respectively). This overall higher DoP for metalinguistic WCF over direct WCF supports similar reports on different Spanish linguistic items (e.g., *ser* vs. *estar* and *preterit* vs. *imperfect*, see Caras, 2019), but also differs substantially from those reports with respect to the percentages of L2 writers processing at a high DoP. Caras (2019) reported that while some of her metalinguistic WCF participants processed at a medium DoP, most in both metalinguistic and direct WCF conditions demonstrated an overall low DoP. In this study, 60% (direct) and

100% (metalinguistic) of the L2 writers demonstrated a high DoP (+H) but remarkably 40% (direct) and 50% (metalinguistic) also processed at the highest level of DoP (+H/+awareness at the level of understanding).

Given that both populations – Caras's (2019) and the current study's – were drawn from the same level of language proficiency in the same language curriculum, several plausible explanations can account for these discrepancies in the DoP data. First, the target linguistic items were different: adjectival agreement and the *gustar* structure versus the copulas *ser* and *estar* and the *preterit* and *imperfect* tenses, which underscores that different types of linguistic items may be processed differentially. Second, the context in which empirical data are gathered may be highly relevant (Leow, 2022). While Caras's participants wrote during their regular class sessions with a time limit of 50-55 minutes, in the current study, writing was performed out of class and there was no time limit. There were several instances of TA protocols lasting over 90 minutes, clearly contributing to the higher DoP observed in this study. Third, while data were gathered on one composition with one topic in Caras's study, the current study collected data from three compositions with different topics that also embedded previously target linguistic items in the requested topic content to address the issue of retention. The opportunity to continue to process the target items in subsequent compositions did appear to have raised the level of processing as exemplified in the overall high DoP coded in the TAs protocols on the second and third compositions. Fourth, while participants in Caras's study were not true beginners (as evidenced on their high pretest scores), those in the current study had minimal high school prior knowledge. This plausibly explains, in addition to the time limit, the difference in DoP between the two studies given that prior knowledge tends to lead to lower levels of cognitive effort (e.g., Bergsleithner, 2019) while highlighting the role of prior knowledge or language proficiency in the WCF strand of research. Fifth, the sample population of this study comprised one intact class versus that of Caras's, which included participants from several sections at the same level. Sixth, participants in this study were taught by one of the researchers, which could have led to a potential Hawthorne effect.[2]

With regard to type of linguistic items (morphological vs. syntactic), the data revealed that processing the noun/adjective agreement with +H DoP/+understanding together with ample practice and exposure appeared to have led to a high level of accuracy by almost all participants. Processing the *me gusta/n* structure (Level 1) with a similar high DoP, even without ample practice or exposure, also appeared

[2] Teachers across all levels who have taken part in the writing project have reported a noticeable improvement in students' performances on both at-home compositions and those produced within a testing condition.

to have produced similar results. However, these L2 writers demonstrated difficulty achieving +DoP/+understanding for the higher level of *gustar* [Level 3: *A Juan le gusta(n)*]; this was captured in the TA protocols that revealed some level of confusion regarding a full understanding of the complex structure. Two plausible explanations can be provided. First, as postulated by Cerezo et al. (2016), the higher level of this syntactic structure is problematic based on the greater number of steps necessary to take to either process or produce these levels. Second, this higher level was introduced late in the semester when compared to the lower level that was introduced earlier. Hence, these L2 writers did not have the amount of exposure to and practice with Level 3 when compared to Level 1. The two L2 writers who did process Level 3 at +H DoP/+understanding also demonstrated accuracy in the final exam. The link between +H DoP/+understanding and subsequent accuracy appears to support similar findings of previous WCF research (e.g., Kim & Bowles, 2019; Park & Kim, 2019) and other strands of research (see Leow, 2019b), including the role of awareness in L2 development (see Leow & Donatelli, 2017 for a recent review).

The data also revealed that while DoP of the target linguistic items remained relatively high during all the rewrites (including the writing of the original compositions) across the semester, the amount of cognitive effort displayed was substantially lower when compared to the first composition (both original and rewrite). Indeed, the speed of activation and appropriateness of knowledge observed in these L2 writers' accuracy of their L2 production on their second and third compositions lend empirical support for the knowledge processing stage of Leow's (2015) model of the L2 learning process in ISLA (see also Bergsleithner, 2019 for similar empirical support) that underlies his feedback processing framework (Leow, 2020). Finally, the beneficial relationship evidenced between high DoP and subsequent L2 learning lends further empirical evidence to several previous studies that have reported a similar type of relationship (see Leow, 2019b for several studies that have reported similar beneficial relationships).

One unique feature of the current study was the participation of one of the researchers as the teacher of the sample population. This participation arguably leads to more insights into participants' performance given the direct knowledge of the classroom setting (format or procedure, structure, type of interaction, etc.), curriculum (textbook, methodology, learning outcomes, homework, test structures, etc.), and the relationships established between the teacher and students during the semester. These are variables all external to the role of WCF that are usually not controlled or even mentioned by previous WCF studies. However, knowing participants can clearly offer insights into their behavioral performances. Participant C was analytic and always highly prepared

for class; he also verbalized the full underlying rules for Levels 1 and 3 of *gustar*, typically processed at the highest level during both composing and revising, and reported several instances of metacognition while composing his compositions (e.g., "I can show my knowledge of time here! Tengo la clase (la clase es femenina), tengo la clase de español a las diez de la mañana"). He wrote 320, 528, and 523 words on his three compositions and spent an average of 90 and 47 minutes on the original compositions and rewrites, respectively. Compared to Participant C, which may raise the potential issue of individual differences in terms of attitude, self-efficacy, and self-confidence, Participant B wrote 267, 367, and 335 words on his three compositions and spent an average of 80 and 26 minutes on his original and rewrites (with more errors), respectively. He claimed around mid-semester in an office meeting that he not only found studying Spanish difficult but also studying in general, wrote in one composition *soy un poco tonto* "I am not smart," and overall demonstrated low self-esteem. While his original compositions revealed much cognitive effort, he demonstrated an overall medium DoP while addressing the WCF during his rewrites. Indeed, as the protocol reported above for this participant revealed, his immediate accuracy on the rewrite, clearly based on addressing the issue of verbal agreement and not as part of the *gustar* structure, did not translate into retention, as measured on the final exam. This processing behavior exemplifies one of the postulations of Leow's (2020) feedback processing framework that explains this immediate performance, typically demonstrated after the provision of feedback but not sustained at a later date.

In addition, the level of accuracy of many of these ten L2 writers on the target linguistic items clearly did not originate solely from the WCF provided on their compositions. Perhaps the WCF could have raised their awareness of their errors but the data appear to indicate that *how* they processed such grammatical information might have played a more important role. Participants also submitted written homework that focused on the target linguistic items covered in the textbook, listened to and read L2 data with the target items, and also practiced them orally (e.g., showing a photo of their family members and describing their characteristics and personalities – agreement), sharing their likes or dislikes (*gustar*) during breakout room sessions or whole class activities.

Perhaps more crucially, when viewed from an ISLA *applied* perspective with curricular and pedagogical ramifications, were two observations, namely, the previous composition format, and the methodology employed in the course. First, the typical composition employed previously in the language curriculum was shorter (e.g., a minimum of 10 lines or 150 words) with a more global composition topic (e.g., "include a description of yourself, place of origin, year and major, and a description of your courses and professors"), which lent itself to open-ended responses that might not have allowed adequate written practice

with recently covered target linguistic items. In the current study, participants/students were guided carefully via the use of specific prompts to address required information that necessitated the use of these items (see Appendix B). This format led students to write more detailed and controlled compositions that, ultimately, provided more focused practice and opportunities for deeper processing of the target items across all compositions at this proficiency level. Second, recall that students were informed early in the semester about the concept of DoP and cognitive engagement and their potential contributions to any type of learning and subsequent retention. Their overall cognitive engagement, well-documented in their preparation for and active participation in all class sessions throughout the semester, may provide a plausible explanation not only for the relatively high use of reactivation of prior linguistic knowledge during the second and third compositions but also for the relatively high level of accuracy (with the exception of Level 3 of gustar) when compared to the relatively low increase of accuracy or correction rates reported in previous WCF studies over a shorter period of time (e.g., Caras, 2019; Shintani & Ellis, 2013). These findings have led to a curricular change in the writing component of the language curriculum, not only at the beginning level but across the entire curriculum.

7. Conclusion, implications, and directions for future research

This preliminary quasi-experimental study sought to add to and advance the process-oriented ISLA research agenda for the *writing-to-learn* WCF strand of research (Leow, 2020; Manchón & Leow, 2020) together with an ISLA *applied* perspective (Leow, 2019a; Leow & Manchón, 2022) that situates the research design within an authentic syllabus and language curriculum. The data revealed, for this small sample, that while an overall high DoP was observed in the TA protocols, the metalinguistic WCF condition demonstrated a higher level of DoP when compared to direct WCF, an unsurprising result. What was notable was the +high DoP/+understanding reported in the TA protocols in both WCF conditions (40% and 50% for direct and metalinguistic WCF, respectively). However, it was also observed that the preferred type of WCF was split evenly by these L2 writers who were exposed to both types during the semester. The link between DoP and performances on related linguistic items on the tests and final exam also appeared to indicate that it may not be the type of WCF L2 writers receive from their teachers that plays an important role in subsequent L2 development but *how* they process such WCF. At the same time, it was also noted that, in a quasi-experimental study embedded within the language curriculum across the entire semester, the role of WCF is arguably not the only variable that may contribute to subsequent L2 development. The writing component is not separate

from other skill components in the curriculum, given that other activities and assignments do include practice with target linguistic items during the course of the semester, and there is repeated but differential exposure to target items in the content of the pedagogical materials throughout the semester. In other words, the methodological or ecological question may be whether we attempt to tease out the role of WCF, quite challenging if adhering to an ISLA *applied* perspective, or simply acknowledge the authentic classroom setting, together with several other variables (e.g., individual differences, curricular differences, etc.) that potentially may impact the pure effect of WCF on subsequent L2 development.

The data also revealed that type of linguistic item (i.e., noun-adjectival agreement vs. two levels of the *gustar* structure) may play a role in how L2 writers process WCF. However, given that types of linguistic items do vary in levels of complexity/difficulty and saliency, in addition to the amount or frequency of exposure and classroom practice, this variable clearly warrants future research. In addition, while this small sample of L2 writers continued to demonstrate a relatively high level of DoP during both the original composing and rewrites of the second and third compositions, the amount of cognitive effort was observed to be reduced when compared to the first composition (both original and rewrite).

Situating the present study within an ISLA *applied* perspective does have major pedagogical and curricular implications. As pointed out in Leow (2020), 1) the data (both online and offline) are authentic as related to the language curriculum (versus a laboratory-based setting), 2) they are gathered longitudinally (versus a one-shot design) within a given syllabus and over the course of a language course, 3) there is ecological validity in the findings, and 4) teachers can easily associate the findings as pertinent to curricular learning outcomes. Two additional pedagogical extrapolations derived from this study may be: 1) the potential value of underscoring the importance of cognitive engagement while composing and revising L2 compositions, which may support the promotion of robust learning from the ISLA *applied writing-to-learn* strand of research (Leow & Manchón, 2022; Manchón & Leow, 2020) and 2) the suggestion to provide students with pertinent topics accompanied by well-designed prompts to promote deeper processing of problematic L2 items. Future WCF studies framed within this process-oriented ISLA *applied* research design are indeed warranted to address, with larger samples, variables that include similar or other types of WCF and linguistic items, level of language proficiency, individual differences, length of compositions, the potential of transfer of modality, different educational levels, etc. together with much needed comparative probing of L2 writers' processing dimensions during both their original and rewrite stages.

References

Bergsleithner, J. M. (2019). The role of prior knowledge in depth of processing during written production: A preliminary investigation. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 104-118). Routledge. https://doi.org/10.4324/9781315165080-33

Caras, A. (2019). Written corrective feedback in compositions and the role of depth of processing. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 188-200). Routledge. https://doi.org/10.4324/9781315165080-33

Cerezo, L., & Caras, A., & Leow, R. P. (2016). Effectiveness of guided induction versus deductive instruction on the development of complex Spanish "gustar" structures: An analysis of learning outcomes and processes. *Studies in Second Language Acquisition*, *38*, 265-291. https://doi.org/10.1017/S0272263116000139

Coyle, Y., Cánovas-Guirao, J., & Roca de Larios, J. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts. *Journal of Second Language Writing*, *42*, 25-43. https://doi.org/10.1016/j.jslw.2018.09.002

Housen, A., & Simoens, H. (2016). Introduction: Cognitive perspectives on difficulty and complexity in L2 acquisition. *Studies in Second Language Acquisition, 38,* 163-175. https://doi.org/10.1017/S0272263116000176

Kim, H. R., & Bowles, M. (2019). How deeply do second language learners process written corrective feedback? Insights gained from think-alouds. *TESOL Quarterly*, *53*(4), 913-938. https://doi.org/10.1002/tesq.522

Leow, R. P. (2012). Explicit and implicit learning in the L2 classroom: What does the research suggest? *The European Journal of Applied Linguistics and TEFL, 2*, 117-129.

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. Routledge. https://doi.org/10.4324/9781315887074

Leow, R. P. (2019a). From SLA > ISLA > ILL: A curricular perspective. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 485-493). Routledge. https://doi.org/10.4324/9781315165080-33

Leow, R. P. (Ed.). (2019b). *The Routledge handbook of second language research in classroom learning*. Routledge. https://doi.org/10.4324/9781315165080-33

Leow, R. P. (2020). L2 writing-to-learn: Theory, research, and a curricular approach. In R. M. Manchón (Ed.), *Writing and language learning: Advancing research agendas* (pp. 95-117). John Benjamins. https://doi.org/10.1075/lllt.56.05leo

Leow, R. P. (2022). The influence of learning contexts on the psycholinguistics of L2 learning. In A. Godfroid & H. Hopp (Eds.), *The Routledge handbook of*

second language acquisition and psycholinguistics (pp. 373-386). Routledge. https://doi.org/10.4324/9781315165080-33

Leow, R. P., & Donatelli, L. (2017). The role of (un)awareness in SLA. *Language Teaching*, *50*, 189-211. https://doi.org/10.1017/S0261444817000039

Leow, R. P., & Manchón, R. M. (2022). Expanding research agendas: Directions for future research agendas on writing, language learning and ISLA. In R. M. Manchón & C. Polio (Eds.), *Routledge handbook of second language acquisition and writing* (pp. 299-311). Routledge. https://doi.org/10.4324/9781315165080-33

Leow, R. P., & Suh, B.-R. (2022). Theoretical perspectives on writing, corrective feedback, and language learning in individual writing conditions. In R. M. Manchón & C. Polio (Eds.), *Routledge handbook of second language acquisition and writing* (pp. 9-21). Routledge.

Manchón, R. M. (Ed.). (2011). *Learning-to-write and writing-to-learn in an additional language*. John Benjamins. https://doi.org/10.1075/lllt.31

Manchón, R. M., & Leow, R. P. (2020). Investigating the language learning potential of L2 writing: Methodological considerations for future research agendas. In R. M. Manchón (Ed.), *Writing and language learning: Advancing research agendas* (pp. 336-355). John Benjamins. https://doi.org/10.1075/lllt.56.14man

Manchón, R. M., Nicolás-Conesa, F., Cerezo, L., & Criado, R. (2020). L2 writers' processing of written corrective feedback: Depth of processing via written languaging. In W. Suzuki & N. Storch (Eds.), *Languaging in language learning and teaching* (pp. 241-265). John Benjamins. https://doi.org/10.1075/lllt.55.11man

Park, E. S., & Kim, O. Y. (2019). Learners' use of indirect written corrective feedback: Depth of processing and self-correction. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 214-228). Routledge. https://doi.org/10.4324/9781315165080-33

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129-158.

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22,* 286-306. https://doi.org/10.1016/j.jslw.2013.03.011

APPENDIX A

Sample of abbreviations and correction symbols sheet (metalinguistic WCF)

A	Personal a	Todavía no he conocido ^ tu novia→Todavía no he conocido a tu novia
ADJ	adjective	Ella es bien estudiante→Ella es buena estudiante Es importa→Es importante
AGR	agreement	El casa blanco > La casa blanca
CONS	construction	Mi hermano y yo nos gustaba...→A mi hermano y a mí nos gustaba...
G	gender	la tema→el tema Él es una persona bueno→Él es una persona buena

APPENDIX B

Composition prompts

Composición 1: *Yo mismo/a y mis clases*
Write a composition about yourself and your classes. Answer each point to guide your composition. [Only pertinent prompts are provided.]

Párrafo (Paragraph) #1:
Introduce yourself including the personal information requested below.

1. Presentación de ti mismo/a (your name and your personal characteristics and personality. You must use at least 2 of the personal characteristic adjectives and 2 of the personality characteristic adjectives provided below PLUS any others if so desired).
   [Características personales: *alto, bajo, mediano, guapo, bonito, delgado, gordo*]
   [Personalidad: *simpático, introvertido, extrovertido, sincero, cómico, divertido, gregario*]
2.-4.

Párrafo #2:

1.
2. Descripción de tus clases y tu estado temporal (describe <u>each class</u> that you are taking and how you are (you feel) in each class. For each class you describe, use adjectives from Clases below to describe the class and adjectives from Estado temporal below to describe your emotional state, PLUS any others if so desired.)
   [Clases: *bueno, aburrido, intensivo, favorito, malo, espantoso, largo (long)*]
   [Estado temporal: *tonto, perdido, serio, interesado, contento, aburrido*]
3.
4. The classes you like
5. The classes you don't like
6. The one class you like the best (*más*)
7. The one class you like least (*menos*)

<u>Puntos de gramática para incluir en la escritura de la composición</u>: artículos; número y género (concordancia); verbos *ser* y estar; los números 0-30; gustar

<u>Extensión:</u> Aproximadamente 300 palabras.

Composición 2: *La información personal y el horario de mis amigos*
Write the personal information and daily schedule (morning, afternoon, and evening) of two friends, one male and the other female. Answer each point to guide your composition. [Only pertinent prompts are provided.]

Párrafo #1:

1. La información de tu compañero y compañera (their names and physical characteristics and personalities. (You must use at least 2 of the personal characteristic adjectives and 2 of the personality characteristic adjectives provided below PLUS any others if so desired).
   [Características: *alto, bajo, mediano, guapo, bonito, delgado, gordo*]

[Personalidad: *simpático, discreto, tímido, creativo, decidido, divertido, metódico*]

2.-7.

8. The classes each one of your friends likes
9. The classes each one doesn't like
10. The class each one likes the most
11. The class each one likes the least

Párrafo #2:

Puntos de gramática:
El número y género (concordancia); verbos *ser* y estar; el tiempo presente, la hora, la frecuencia (siempre, nunca, a veces etc.), gustar

Extensión: Aproximadamente 300 palabras.

APPENDIX C

TA practice instruction

Solve the following math word problem by thinking your thoughts aloud while the audio recording app is recording your voice. "Thinking your thoughts aloud" is like talking to yourself aloud or, literally, saying aloud whatever is going through your head at that moment. You may speak in whatever language you prefer, but please be sure to speak in a loud, clear voice so that the audio recording app can pick up your voice. Please note that the purpose of this math word problem is for you to practice thinking your thoughts aloud, not to have your math skills assessed.

Math Word Problem: Theodore went to the supermarket. He bought two dozen eggs for $2.50/dozen, organic milk for $5.49, laundry detergent for $12, and eight cartons of yogurt at 80 cents each. How much did Theodore spend?

APPENDIX D

Participant questionnaire

1. Your instructor used two different types of feedback methods listed below this semester. Which composition feedback do you believe was more beneficial to the development of your Spanish writing? [Options: Direct feedback (instructor corrects your errors by providing the correct wording), Metalinguistic feedback (instructor writes correction symbols and you figure out how to correct your error]
2. Why do you think the method you chose helped you improve more than the other method?
3. If you spoke your thoughts aloud while writing, what did you think of that experience?